(2015)

The Stretchable Arms for Collaborative Remote Guiding

Morgan Le Chénéchal[†1], Thierry Duval[‡1,2], Valérie Gouranton[§1,3], Jérome Royan[†1] and Bruno Arnaldi[§1,3]

[1]IRT b<>com, France
[2]Télécom Bretagne / LABSTICC, France
[3]INSA Rennes / IRISA/Inria, France

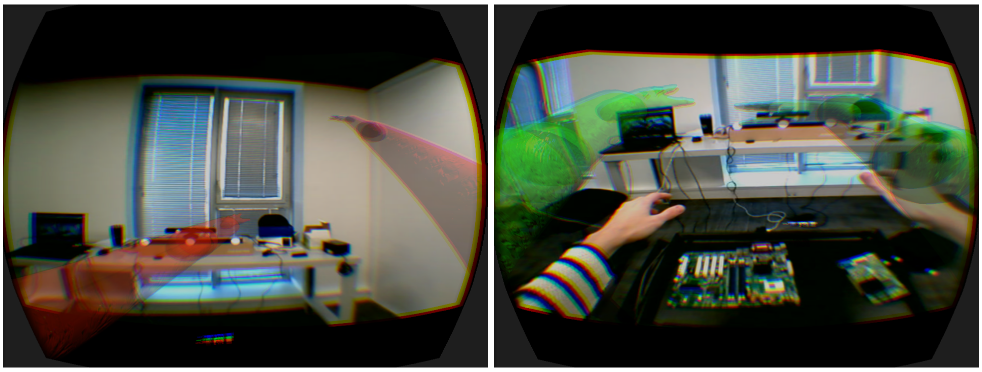


Figure 1: The stretchable guiding arms, controlled by the remote expert's location, from the agent's viewpoint. On the left, the arms are extended and red because the guide is ahead of the agent. On the right, the agent has reached the guide's location and the guiding arms retrieve their initial length and green color.

Abstract

The help of a remote expert to guide an agent in performing a physical task can be advantageous in many ways: saving time and money by avoiding travel, and thus increasing the rate of intervention. In many situations, the remote expert wishes to guide the agent by first placing him in the correct location to achieve the task. However, as the agent is not a robot, the expert can not use a location controller to place the agent. Instead, interaction techniques must enable the expert to achieve this task before physical manipulation guidance. In this paper, we propose a novel interaction technique for remote guiding based on arm gestures. First, the remote expert (using a VR setup) virtually collocates himself with the agent (using an AR setup), then controls virtual arms collocated with both users' shoulders. Second, if the expert starts to move forward to grasp a virtual object, the virtual arms start to stretch in order to keep the shoulders' collocation on the agent's side. This metaphor allows the agent to understand the direction of the expert's motion easily while preserving the naturalness of the interaction and avoiding the use of a frustum to represent the expert's head location.

Categories and Subject Descriptors (according to ACM CCS): Information Interfaces and Presentation (e.g. HCI) [H.5.2]: User Interfaces—Computer Graphics [I.3.6]: Methodology and Techniques—

1. Introduction

Collaborative Virtual Environments (CVE) allow multiple users to share a common virtual world to achieve collaborative tasks. For a long time, the literature leveraged the issue to get the user to reach a specific viewpoint in order to show him some artifact. Indeed, in the real world, natural

† {morgan.lechenechal, jerome.royan}@b-com.com
‡ thierry.duval@telecom-bretagne.eu
§ {valerie.gouranton, bruno.arnaldi}@irisa.fr

awareness features such as gaze direction [SWM*08] allows people to communicate without the need to explain explicitly but these are rarely available in Mixed Reality (MR) applications. Thus, different techniques must be proposed to handle this useful feature: **to guide someone to reach a specific viewpoint**.

We propose a novel interaction technique based on a MR system in which a remote expert using a Virtual Reality (VR) application guides an agent using an Augmented Reality (AR) application. The agent must perform a physical task that is guided by virtual arms collocated with his shoulders, controlled by the remote expert. Before performing the actual arm guidance, the remote expert must reach the correct location in the virtual world while showing it to the agent in his augmented real world. To perform this initial location guidance, we propose to extend the GoGo technique initially developed for single user selection and manipulation tasks in VR [PBWI96]. We added the arms representation as in [FBHH99] and removed the non-linear scaling, instead controlling it directly with the expert's interactions.

Section 2 introduces proposed techniques in the literature that have achieved a similar goal and their limitations. Section 3 presents our stretchable guiding arms technique as well as its context of use, and Section 4 discusses the results of a pilot user study. Finally, Section 5 concludes this paper.

2. Related Work

Making a user reach a specific viewpoint is not only used in CVE. Indeed, many single user MR applications need to provide such a feature. Recently, Sukan et al. proposed the para-frustum approach to guide a user in an AR application to reach a set of acceptable viewpoints [SEO*14]. It presents an advanced technique based on a 3D augmentation that shows the user the viewpoint he must reach. Compared to classical approaches such as the use of head-up display (HUD), this technique greatly eases the reaching process because of a fully collocated guidance in the real interaction space. However, Sukan et al. highlight the complexity of defining the shape of the 3D augmentation, and thus it seems complicated to extend it for collaborative purposes with an online definition of the shape by one user to another.

In CVE, we distinguish two approaches: asynchronous and synchronous guidance. First, asynchronous guidance is often used when a user defines some interesting viewpoint to another user, who can switch from one predefined viewpoint to another. In [DFNA08], these interesting viewpoints are represented with virtual cameras all around a set of scientific data visualizations. Second, synchronous guidance is the category to which our technique belongs. It is utilized by a user who interactively defines a viewpoint for another user to reach. Many metaphors have been proposed [NDF13]: directional arrows, compass, lit path. However, these techniques work well in an asymmetric setup, i.e. when the guide has a global view of the scene and while the visitor is immersed in it, but are hardly extendable for a symmetric, fully collocated setup. Indeed, even if their implementations can be fully automated, i.e. the guide does not need to define the parameters of the guidance cues himself because the system is able to compute them automatically, they suffer from strict limitations: the loss of naturalness in the guidance process for the guided user because of the interpretation of specific visual cues and the eye strain that can be generated by the proximity of visual cues according to the guided viewpoint.

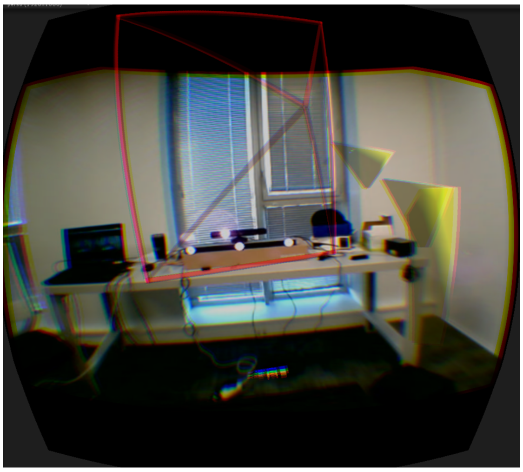


Figure 2: A simple guiding technique based on directional arrows and viewing frustum of the remote expert.

3. The Stretchable Guiding Arms Technique

The stretchable guiding arms technique has been developed to overcome the limitations found in the techniques from the literature in a specific context: the expert and the agent share a fully collocated setup. Moreover, the final goal of the system is to guide the agent in performing a physical task with the use of virtual arms controlled by the remote expert. In this setting, the first step consists of viewpoints collocation. Then, the expert must show which object to move, and thus he often needs to move the agent's viewpoint forward in order to grasp the virtual cloned object. The existing technique based on directional arrows can achieve this purpose, as illustrated in Figure 2, but first users who tested our system complained about eye strain because of the proximity of the visual cues, including the remote expert's frustum. Moreover, it increases the cognitive load on the agent when having to follow arms guidance and location guidance at the same time using different visual cues. To avoid this, we take advantage of the already existing virtual arms initially present for gesture-based guiding purposes only. They can also be used for location reaching guidance in a non-intrusive way, contrary to other methods. Figure 3 illustrates the principle of the proposed stretchable guiding arms technique. When

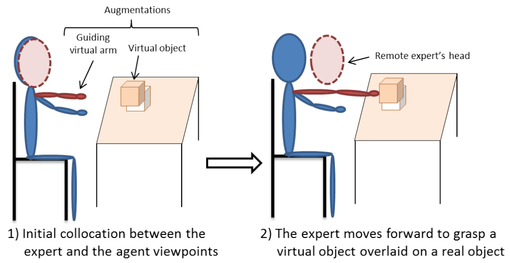


Figure 3: Principle of the stretchable guiding arms technique. Illustration from the agent side in blue. The red circle represents the expert's head but is not actually displayed. Here, it is added to understand the expert's interaction. From the expert's viewpoint, his virtual arms do not scale and also stay collocated with his shoulders.

the expert chooses to switch to stretch mode, his forward motion makes the virtual arms seen from the agent's viewpoint start to scale up to keep the virtual hands always collocated with the expert's ones, while maintaining the virtual shoulders fixed according to the agent's shoulders. From the expert's viewpoint, his virtual arms do not scale up, and also stay collocated with his shoulders.

The main advantage of this technique is the naturalness of the guidance. Indeed, the agent can directly perceive the forward motion of the remote expert, thanks to the scaling up of his virtual arms. Then, to follow the expert, the agent simply moves his viewpoint in order to retrieve the initial length of his virtual arms. In addition to the scaling cue, a coloring code smoothly changes the color of the independent virtual arms from green to red, according to their scale. They return to green when the scale is close to its initial value, which means that the agent has reached the expert's viewpoint (cf. Figure 1). If the agent overtakes the expert's location, the virtual arms keep their initial length and color, but their bending indicates to the agent that he must move backward to retrieve a correct pose.

In parallel, the expert can use a virtual camera with a different viewpoint of his VE in order to get visual feedback on the agent following. This camera is displayed in a corner of the expert's viewport and is enabled when the stretchable guiding arms technique is activated. The remote expert can switch between different available viewpoints: a side or front view based on his head location, or the viewpoint of the agent slightly behind him when performing a forward motion guidance. Moreover, the color change for virtual arms from green to red is used to inform the remote expert explicitly about his distance with the guided agent.

3.1. Implementation

Currently available see-through AR glasses do not provide a wide enough field of view. Thus, we use a see-through HMD equipped with a pair of stereo cameras. The prototype has been developed with Unity3D™to be used with an Oculus Rift™DK2. On the expert side, the virtual arms are controlled with a Razer™Hydra that provides two handles with 6DOF tracking and joysticks for navigation.

4. Pilot User Study

We compared reaching performances and subjective feelings of (C1) our technique (cf Figure 1) with (C2) the frustum with 3D directional arrows based approach (cf. Figure 2). In order to evaluate the agent's performance in the location reaching process, we try to remove bias generated by the expert's interactions. Thus, we recorded the interactions of a real expert in a pre-process phase in order to replay it for the experiment, removing possible expert's interaction bias due to changing behavior. We postulate that the expert's interactions are not really affected by the agent's ones. Indeed, the expert must reach a location to grab a virtual object, and thus must move forward to achieve the goal assuming that the agent is following.

4.1. Measurements and Hypotheses

We collected three objective measures: completion time (hit target around 0.15m), mean distance between the replayed expert and the subject, and final precision (triggered by the subject) in location and orientation. Moreover, subjects fill in a final subjective questionnaire (7 point Likert-scale) to express their impressions of the naturalness of the system, their perceived cognitive load and eye-strain, amongst other considerations. We stated the following hypotheses:

- H1: Completion time, mean distance and precision are equivalent in both conditions C1 and C2
- H2: C1 is more comfortable to use than C2

4.2. Protocol

We design our experiment in a counter-balanced way using a set of 20 pre-defined targets (10 per condition) in a range from 1 to 3 meters from the home location and a horizontal angle range from -30 to 30 degrees. Their altitude is defined around a mean value of 1.5m with a range of ± 0.3m. Each participant passes both conditions twice alternatively, with a different set of 2x5 virtual targets per condition. These targets are not displayed for the replay process because the subjects must reach them following only the recorded guide. We simulate the initial viewpoints collocation, illustrated in step 1 of the process described in Figure 3, with a re-targeting of the replayed data according to the subject's height and exact home location.

4.3. Results

Results have been collected from 10 subjects aged from 20 to 30 ($mean = 24.4, sd = 4.03$). A few iterations are not part

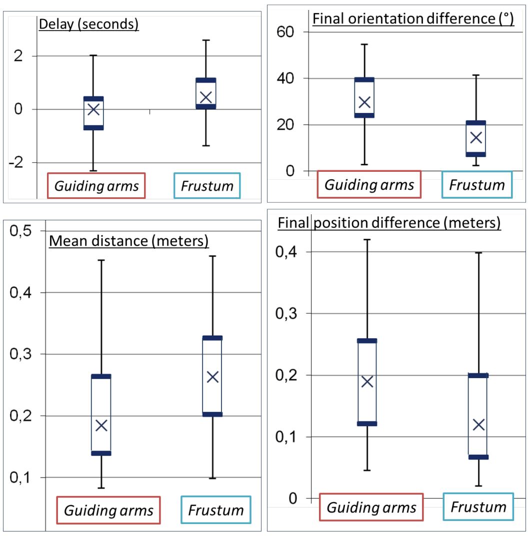


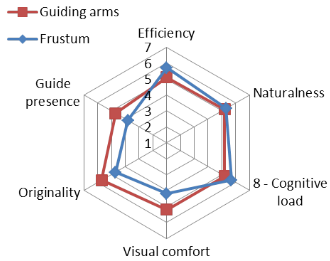
Figure 4: Boxplots of the delay between the target hitting by the guide and the subject, their mean distance until the hit, and their final position and orientation differences.

of the analysis because of tracking issues. Figure 4 illustrates the collected quantitative results. We found a significant difference in the target hitting delay between the guide and the subject ($F(1,194) = 20.47, p < .01$). This delay is shorter with (C1) the guiding arms ($M = -0.03, sd = 1.28$) than with (C2) the frustum ($M = 0.75, sd = 1.61$). In the same way, the mean distance between the guide and the subject until the hit is significantly smaller ($F(1,194) = 20.39, p < .01$) using (C1) the guiding arms ($M = 0.21, sd = 0.008$) than (C2) the frustum ($M = 0.26, sd = 0.007$). In terms of precision, the frustum-based guiding (C2) is significantly more accurate than the guiding arms (C1) for both location ($F(1,194) = 11.34, p < .01$; (C1): $M = 0.19, sd = 0.006$; (C2): $M = 0.15, sd = 0.01$) and orientation ($F(1,194) = 55.37, p < .01$; (C1): $M = 31.10, sd = 122.55$; (C2): $M = 16.54, sd = 252.88$).

4.4. Discussion

H1 is not proved. Actually, we found that the guiding arms (C1) are better for closely following the remote guide, while the frustum-based guiding (C2) is more accurate for final precision. This degradation of preciseness using C1 can be explained by the complexity of human body's DOFs that generate variations between the relative pose of the agent's head and the expert's hands. Thus, the coupling of C1 with an alternative guiding technique for final positioning should be studied to improve preciseness without decreasing naturalness of the guidance. Concerning the qualitative results (cf. Figure 5), we validated H2 because of better visual comfort using our approach (C1). Moreover, subjects prefer our approach (C1) in terms of originality and guide presence. For other criterion (efficiency, naturalness and cognitive load), results are similar for both conditions C1 and C2.

Figure 5: Synthesis of subjects' qualitative appreciations.



5. Conclusion

We presented an extension of the GoGo interaction technique for collaborative purposes. We use the same idea of virtual arms scaling up, but not in a selection and manipulation context. Rather, we use it to guide a remote user to reach a location defined by the location of another one. Solutions in the literature require the use of specific visual cues that decrease the naturalness as well as the immersion of the application. Our stretchable guiding arms technique takes advantage of already existing virtual arms that are used to guide a physical task using remote arm gestures. We believe our technique is more appropriate in this context, and the results of the pilot user study show that following performances are even better using our approach as well as enhancing comfort of use. Future work will focus on improving final precision of our guiding technique.

References

[DFNA08] DUVAL T., FLEURY C., NOUAILHAS B., AGUERRECHE L.: Collaborative exploration of 3d scientific data. In *Proceedings of the 2008 ACM symposium on VRST* (2008), ACM, pp. 303–304. 2

[FBHH99] FRASER M., BENFORD S., HINDMARSH J., HEATH C.: Supporting awareness and interaction through collaborative virtual interfaces. In *Proceedings of the 12th annual ACM symposium on UIST* (1999), pp. 27–36. 2

[NDF13] NGUYEN T. T. H., DUVAL T., FLEURY C.: Guiding Techniques for Collaborative Exploration in Multi-Scale Shared Virtual Environments. In *Proceedings of GRAPP* (2013). 2

[PBWI96] POUPYREV I., BILLINGHURST M., WEGHORST S., ICHIKAWA T.: The go-go interaction technique: non-linear mapping for direct manipulation in VR. In *Proceedings of the 9th annual ACM symposium on UIST* (1996), ACM, pp. 79–80. 2

[SEO*14] SUKAN M., ELVEZIO C., ODA O., FEINER S., TVERSKY B.: ParaFrustum: visualization techniques for guiding a user to a constrained set of viewing positions and orientations. In *Proceedings of the 27th annual ACM symposium on UIST* (2014), ACM, pp. 331–340. 2

[SWM*08] STEPTOE W., WOLFF R., MURGIA A., GUIMARAES E., RAE J., SHARKEY P., ROBERTS D., STEED A.: Eye-tracking for avatar eye-gaze and interactional analysis in immersive collaborative virtual environments. In *Proceedings of the 2008 ACM conference on CSCW* (2008), pp. 197–200. 2